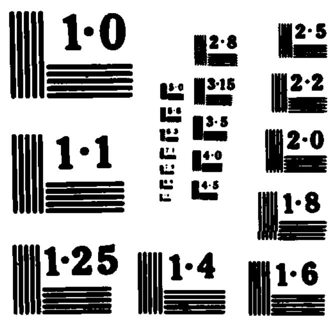
AD-A186 722 INFERENCE FOR THE EXPONENTIAL LIFE DISTRIBUTION(U) 1/1
FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS
R E BARLOW ET AL. 1985 FSU-STATISTICS-M-368-B

UNCLASSIFIED AFOSR-TR-87-1013 $AFOSR-77-3179 F/G 12/1

END
DATE
FILMED
1 88
DTIC


1·0
2·8
2·5
3·15
2·2
3·5
2·0
4·0
1·1
4·5
1·8
1·25
1·4
1·6

AD-A186 722

UNCLASSIFIED DTIC FILE COPY

SECURITY CLASSIFICATION OF THIS PAGE

(1)

DTIC ELECTE OCT 1 6 1987 S D

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY: NA	3. DISTRIBUTION/AVAILABILITY OF REPORT
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE: NA	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): M-568-R - P-463	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 87-1013 AFOSR 85-174
6a. NAME OF PERFORMING ORGANIZATION: FLORIDA STATE UNIVERSITY	6b. OFFICE SYMBOL (If applicable)
7a. NAME OF MONITORING ORGANIZATION: AFOSR	
6c. ADDRESS (City, State and ZIP Code): TALLAHASSEE, FL 32306-3033	7b. ADDRESS (City, State and ZIP Code): Bldg. 410 Bolling AFB, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR; 8b. OFFICE SYMBOL (If applicable): NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-77-3179
8c. ADDRESS (City, State and ZIP Code): BLDG. 410 BOLLING AFB, DC	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: 61102F; PROJECT NO.: 2304; TASK NO.: A5; WORK UNIT NO.:
11. TITLE (Include Security Classification): Inference for the Exponential Life Distribution	
12. PERSONAL AUTHOR(S): Barlow, R.E., Proschan, Frank	
13a. TYPE OF REPORT: ~~REPRING~~ Reprint; 13b. TIME COVERED FROM ____ TO ____	14. DATE OF REPORT (Yr., Mo., Day): 04/16/86; 15. PAGE COUNT: 143 - 164
16. SUPPLEMENTARY NOTATION	
17. COSATI CODES: FIELD / GROUP / SUB. GR.	18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Our objective is to develop methodology for analyzing life test data. Initially, ~~we have~~ only data—no mathematical models. Through an exploratory data analysis or an analysis based on the physical processes generating the data, ~~we may judge~~ an exponential life distribution model as appropriate for the analysis of the data. Specifically:

$$F(x|\lambda) = 1 - \exp[-\lambda x], \qquad x > 0,\ \lambda > 0, \tag{0.1}$$

where λ is the unknown constant failure rate. The vertical bar in $F(x|\lambda)$ indicates that we are conditioning on the parameter λ; *i.e.*, for specified λ the distribution is exponential with failure rate λ. The corresponding density is

$$f(x|\lambda) = \lambda \exp[-\lambda x], \qquad x > 0,\ \lambda > 0. \tag{0.2}$$

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: ~~FRANK PROSCHAN~~ Maj Crowley	22b. TELEPHONE NUMBER (Include Area Code): ~~644-3218~~; 22c. OFFICE SYMBOL: NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

SECURITY CLASSIFICATION OF THIS PAGE

AFOSR-TR- 87-1013

Inference for the Exponential Life Distribution.

R. E. BARLOW

Department of Industrial Engineering and Operations Research
University of California - Berkeley, CA 94720

F. PROSCHAN

Department of Statistics, Florida State University - Tallahassee, FL

Our objective is to develop methodology for analyzing life test data. Initially, we have only data—no mathematical models. Through an exploratory data analysis or an analysis based on the physical processes generating the data, we may judge an exponential life distribution model as appropriate for the analysis of the data. Specifically:

$$F(x|\lambda) = 1 - \exp[-\lambda x], \qquad x \geqslant 0, \ \lambda > 0, \tag{0.1}$$

where λ is the unknown constant failure rate. The vertical bar in $F(x|\lambda)$ indicates that we are conditioning on the parameter λ; *i.e.*, for specified λ the distribution is exponential with failure rate λ. The corresponding density is

$$f(x|\lambda) = \lambda \exp[-\lambda x], \qquad x \geqslant 0, \ \lambda > 0. \tag{0.2}$$

1. – Basic concepts.

To begin with, we suppose the life test data consist of observed complete lifetimes $x_1, x_2, \ldots, x_n$ on n units. For example, table 1.I lists lifelengths ordered by rank of 100 Kevlar 49/Epoxy strands subjected to a high static load [1]. An exploratory data analysis indicates that an exponential life distribution may be appropriate for analyzing these lifelengths. Thus we assume that the observations constitute a sample of n independent, identically distributed random variables with distribution F given by (0.1). Although we assume a fixed λ exists which specifies F, we are uncertain as to the true value of λ and seek a method which uses the data to express probabilistically our uncertainty regarding the true value of λ.

TABLE 1.I. – *Times to failure of strands subjected to stress at 80 % of mean rupture strength.*

Rank	Hours	Rank	Hours	Rank	Hours	Rank	Hours
1	1.8	26	84.2	51	152.2	76	285.9
2	3.1	27	87.1	52	152.8	77	292.6
3	4.2	28	87.3	53	157.7	78	295.1
4	6.0	29	93.2	54	160.0	79	301.1
5	7.5	30	103.4	55	163.6	80	304.3
6	8.2	31	104.6	56	166.9	81	316.8
7	8.5	32	105.5	57	170.5	82	329.8
8	10.3	33	108.8	58	174.9	83	334.1
9	10.6	34	112.6	59	177.7	84	346.2
10	24.2	35	116.8	60	179.2	85	351.2
11	29.6	36	118.0	61	183.6	86	353.3
12	31.7	37	122.3	62	183.8	87	369.3
13	41.9	38	123.5	63	194.3	88	372.3
14	44.1	39	124.4	64	195.1	89	381.3
15	49.5	40	125.4	65	195.3	90	393.5
16	50.1	41	129.5	66	202.6	91	451.3
17	59.7	42	130.4	67	220.2	92	461.5
18	61.7	43	131.6	68	221.3	93	574.2
19	64.4	44	132.8	69	227.2	94	653.3
20	69.7	45	133.8	70	251.0	95	663.0
21	70.0	46	137.0	71	266.5	96	669.8
22	77.8	47	140.2	72	267.9	97	739.7
23	80.5	48	140.9	73	269.2	98	759.6
24	82.3	49	148.5	74	270.4	99	894.7
25	83.5	50	149.2	75	272.5	100	974.9

The first step is to evaluate the joint-probability density of the random lifetimes $X_1, \ldots, X_n$ evaluated at the observed values $x_1, \ldots, x_n$. Since we are assuming that $X_1, \ldots, X_n$ are independent given λ, the joint density of the observed values is

$$\prod_{i=1}^{n} f(x_i|\lambda) = \lambda^n \exp\left[-\lambda \sum_{i=1}^{n} x_i\right]. \tag{1.1}$$

1·1. ***The likelihood function.*** – To focus attention on the parameter of interest λ, we regard (1.1) as a function of λ and call

$$L(\lambda | x_1, \dots, x_n) = \lambda^n \exp\left[-\lambda \sum_{i=1}^{n} x_i\right] \tag{1.2}$$

the *likelihood function*. (The likelihood, although a function of the parameter λ, is *not* a probability density in the parameter. Hence the vertical bar in L is used to indicate that the data to the right of the vertical bar are given.) The likelihood function provides a means of quantifying the information contained in the data concerning the unknown true value of the exponential parameter λ.

Suppose a unique value $\hat{\lambda}$ of λ exists maximizing the likelihood function. then we call $\hat{\lambda}$ the *mode* of $L(\lambda | x_1, \dots, x_n)$ and the *maximum-likelihood estimator* (MLE) of λ. In general, the MLE, when it exists, is a very useful concept.

To simplify the calculation of $\hat{\lambda}$, we use the fact that the maximum of the likelihood, when it exists, is achieved at the same value of λ as is the maximum of the logarithm of the likelihood. Thus we compute

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \ln L(\lambda | x_1, \dots, x_n) = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i$$

and set the derivative equal to 0. We readily obtain

$$\hat{\lambda} = n\left(\sum_{i=1}^{n} x_i\right)^{-1}$$

and verify that $\hat{\lambda}$ maximizes $L(\lambda | x_1, \dots, x_n)$ for fixed $x_1, \dots, x_n$.

The MLE $\hat{\lambda}$ may be a very satisfactory estimator of the unknown failure rate λ for moderate to large sample sizes n. (Caution: For more complex life distribution models involving an infinite number of unknown parameters, the MLE of the parameters may be quite misleading. See [2] for an example in which the MLE converges to the *wrong* set of the parameter values even though the sample size tends to infinity. Also see [3], p. 12, for a similar two-parameter example and [4], p. 34, for a one-parameter example.) In the present case of estimation of the single parameter λ of the exponential, our uncertainty as to the true value of λ stems from the fact that our sample size n is finite.

We express our uncertainty concerning λ by means of a ***probability distribution*** for λ. To display explicitly this point of view, we let $\tilde{\lambda}$ denote a random variable expressing our uncertainty concerning the unknown true value of λ.

Bayes' theorem. A key theorem based on this point of view is the fundamental Bayes' theorem. It provides a method for computing the probability density of the random variable expressing our uncertainty concerning the parameter conditioned on the observed data.

1.2. *Theorem* (*Bayes' theorem*). – Let *a*) X and $\tilde{\theta}$ be random variables with joint-probability density $p(x, \theta)$, *b*) $p(x|\theta)$ and $p(\theta|x)$ denote the corresponding conditional densities, and *c*) $\pi(\theta)$ denote the marginal density of $\tilde{\theta}$. Let Θ be the parameter space, *i.e.* $\theta \in \Theta$. Then

$$p(\theta|x) = \frac{p(x|\theta)\pi(\theta)}{\int_{\Theta} p(x|\theta)\pi(\theta)\,\mathrm{d}\theta}\,. \tag{1.3}$$

Proof. The joint density $p(x, \theta)$ of X and $\tilde{\theta}$ may be written as

$$p(x, \theta) = p(x|\theta)\pi(\theta)\,.$$

By definition of a conditional probability density,

$$p(\theta|x) = p(x, \theta)/p(x)\,,$$

when $p(x) > 0$, where

$$p(x) \stackrel{\text{def}}{=} \int_{\Theta} p(x|\theta)\pi(\theta)\,\mathrm{d}\theta\,.$$

By combining the three equalities in the steps just above, we immediately obtain the desired conclusion (1.3). ||

Prior and posterior distributions. Before analyzing statistical data, it is helpful and efficient to assess prior knowledge. A convenient way to accomplish this is to formulate a probability density on the parameter(s) of the model selected. Once we select an appropriate model, and a prior distribution on the parameter space for that model, we may complete a useful and informative data analysis in an unambiguous fashion using only the standard calculus of probability theory.

The prior density. First, we confine our choice of prior densities to *proper* densities. A density $\pi(\cdot)$ is proper if $\int \pi(\lambda)\,\mathrm{d}\lambda$ exists and equals one.

Next, to motivate the concept of a natural conjugate prior for $\tilde{\lambda}$, we suppose that, in the particular problem under discussion, we have very little prior information concerning $\tilde{\lambda}$. It seems natural to assume initially a rectangular prior density:

$$\pi_0(\lambda) = \begin{cases} M^{-1} & \text{for } 0 < \lambda < M\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

where M is a very large number (say $M = 10^{12}$). Under this assumption on the prior density, we assign the same probability that $\tilde{\lambda}$ is in any interval in $[0, M]$

of a specified length. For example, the *a priori* probability that $5 \leqslant \tilde{\lambda} \leqslant 10$ is the same as the *a priori* probability that $19 \leqslant \tilde{\lambda} \leqslant 24$.

The posterior density. Suppose we have observed a sample of n lifelengths $x_1, \ldots, x_n$ having joint density $p(x_1, \ldots, x_n|\lambda)$. By Bayes' theorem, the *posterior density* of $\tilde{\lambda}$ based on $\pi_0(\lambda)$ is given by

$$\pi_1(\lambda|x_1, \ldots, x_n) = p(x_1, \ldots, x_n|\lambda)\pi_0(\lambda)\Big/\int_0^\infty p(x_1, \ldots, x_n|\lambda)\pi_0(\lambda)\,d\lambda\,.$$

Recall that the *likelihood* $L(\lambda|x_1, \ldots, x_n) = p(x_1, \ldots, x_n|\lambda)$, namely the probability density of the observed outcome considered as a function of the parameter λ. Thus we may write

$$\pi_1(\lambda|x_1, \ldots, x_n) = L(\lambda|x_1, \ldots, x_n)\pi_0(\lambda)\Big/\int_0^\infty L(\lambda|x_1, \ldots, x_n)\pi_0(\lambda)\,d\lambda\,.$$

Since λ has been integrated out in the denominator, the denominator is now a constant with respect to λ. Hence

$$\pi_1(\lambda|x_1, \ldots, x_n) \propto L(\lambda|x_1, \ldots, x_n)\pi_0(\lambda)\,,$$

where $\propto$ means « proportional to ». Thus the right-hand side is the same as the left-hand side up to a constant which does not depend on the parameter λ. Notice that, since the data $x_1, \ldots, x_n$ have already been observed, the data are *not* considered variables at this stage of the analysis.

Assuming the rectangular prior $\pi_0(\lambda) = M^{-1}$ for $0 \leqslant \lambda \leqslant M$, we obtain for the posterior density of $\tilde{\lambda}$

$$\pi_1(\lambda|x_1, \ldots, x_n) = M^{-1}\lambda^n \exp\left[-\lambda \sum_1^n x_i\right]\Big/M^{-1}\int_0^M \lambda^n \exp\left[-\lambda \sum_1^n x_i\right]d\lambda\,.$$

For $n^{-1}\sum_1^n x_i \gg M^{-1}$, this is approximately

$$\pi_1(\lambda|x_1, \ldots, x_n) \doteq \left(\sum_1^n x_i\right)^{n+1}\lambda^n \exp\left[-\lambda \sum_1^n x_i\right]\Big/\Gamma(n+1)\,, \tag{1.4}$$

where $\Gamma(n+1) = \int_0^\infty u^n \exp[-u]\,du$ is the $n!$ function. In computing (1.4) we have used the fact that $\int_0^\infty c^{n+1}u^n \exp[-cu]\,du = \Gamma(n+1)$ for all $c > 0$.

Thus, if we assume initially a rectangular prior on $[0, M]$, M large, the resulting posterior density is approximately of the form

$$b^a\lambda^{a-1}\exp[-b\lambda]/\Gamma(a) \qquad \text{for } \lambda > 0, \tag{1.5}$$

where $a,\ b > 0$. This is approximately a gamma density with shape parameter $a = n + 1$ and scale parameter

$$b = \sum_{i=1}^{n} x_i .$$

Now, suppose we obtain an additional independent random sample of lifelengths $y_1, y_2, \ldots, y_m$. Then it is reasonable to use as our *new* prior density the posterior density (1.4) obtained from the previous sample. Using as our new prior

$$\pi_1(\lambda) = b^a \lambda^{a-1} \exp[-b\lambda]/\Gamma(a)$$

with $a = n + 1$ and $b = \sum_1^n x_i$, we obtain

$$\pi_2(\lambda|y_1, \ldots, y_m) = \frac{\lambda^m \exp\left[-\lambda \sum_1^m y_i\right] b^a \lambda^{a-1} \exp[-b\lambda]/\Gamma(a)}{\int_0^\infty \lambda^m \exp\left[-\lambda \sum_1^m y_i\right][b^a \lambda^{a-1} \exp[-b\lambda]/\Gamma(a)]\, d\lambda},$$

or

$$\pi_2(\lambda|y_1, \ldots, y_m) = \left(b + \sum_1^m y_i\right)^{m+a} \lambda^{m+a-1} \exp\left[-\lambda\left(b + \sum_1^m y_i\right)\right]\Big/\Gamma(m+a) . \tag{1.6}$$

Thus a is increased by the additional number m of observed failures and b is increased by the additional quantity $\sum_1^m y_i$ to obtain the new posterior density π_2; note, however, that the *form* of the posterior density, the gamma, is retained.

Because of this preservation property (the gamma prior used in the exponential model leads to a gamma posterior), the gamma prior is called the « natural conjugate » prior for the parameter λ in the exponential model. More generally, a family of prior distributions is « conjugate » with respect to a given statistical model if the form of the posterior in each case is the same as that of the prior and it is the minimal such family; parameters of the posterior distributions will, of course, change in accordance with the data observed.

In the present case, we can interpret the prior density parameter $a - 1$ (if a is an integer > 1) as the number of observations in a previous experiment (actual or conceptual) and b as the corresponding total time on test.

In the present exponential model, the gamma prior for λ is mathematically convenient and has intuitive interpretation in terms of an equivalent sample. However, the analyst is not confined to a choice within this family. Rather, the choice of the prior distribution should always reflect the best possible specification of the analyst's prior information concerning the unknown parameter. Thus, in reporting the results of the data analysis, the analyst should present his specification of the prior and the basis for his choice.

1'3. *Example.* – Table 1.I lists the observed lifetimes of 100 organic fiber strands subjected to a static load of 8009 g which corresponds to 80% of their mean rupture strength. Experience has shown that the lifetime of an organic fiber strand at relatively high stress can be reasonably well fitted by an exponential life distribution. Thus we assume

$$P[\text{lifelength} > x] = \exp[-\lambda x] \qquad \text{for } x \geqslant 0,$$

where λ is unknown. As described above, we calculated the MLE of λ as

$$\hat{\lambda} = 4.78 \cdot 10^{-3}/\text{hour}.$$

Since the sample size of 100 is moderately large, the likelihood

$$L(\lambda | x_1, \dots, x_{100}) = \lambda^{100} \exp\left[-\lambda \sum_{1}^{100} x_i\right]$$

will override in importance the retangular prior

$$\pi_0(\lambda) = M^{-1}, \qquad 0 < \lambda \leqslant M,$$

when $M \gg \hat{\lambda}$. From the prior π_0 we calculate the posterior density of $\tilde{\lambda}$ to be approximately a gamma with parameters $a = 101$ and $b = 20.917$ hours. (Note that a is dimensionless while b is measured in hours). See fig. 1.1. The mode of the posterior density may be used to estimate λ.

Sufficient statistics. By a *statistic*, we mean a function (possibly vector-valued) of the data. A statistic, of course, is often used to estimate an unknown parameter of interest. Clearly, the complete set of data observed is trivially a statistic. For a large set of data, working with *all* of the observations may be tedious or even unmanageable. Thus we are motivated to find a statistic of smaller dimensionality like the sample mean (of dimension one) or the number of failures *and* the total time on test (of dimension two), but which contains all of the information in the sample concerning the parameter. Preferably, we would like to estimate the parameter using a statistic of lower dimensionality which summarizes all of the information in the sample. We will often use D to denote the observed data. For example, D could stand for the vector of observed values $(x_1, x_2, \dots, x_n)$. Later it will denote more complicated data sets. This motivates

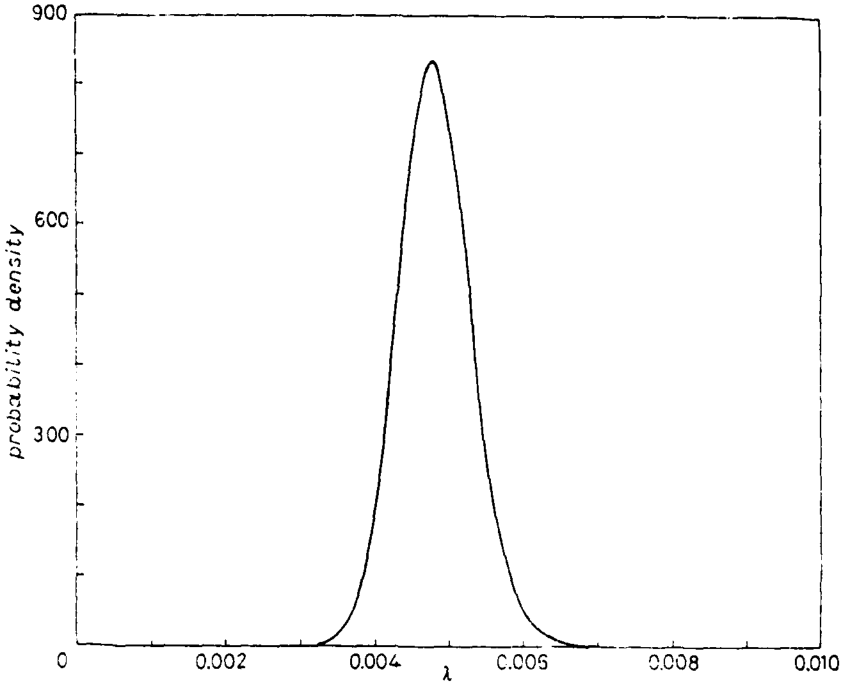


Fig. 1.1. – Computer plot of posterior gamma density of lambda.

1'4. *Definition.* – Let D denote the data with probability density $p(D|\theta)$ indexed by the parameter θ. A statistic s is *sufficient* for θ if and only if, for *every prior* $\pi(\theta)$, the posterior

$$\pi(\theta|D) = p(D|\theta)\pi(\theta)\Big/\int_{\Theta} p(D|\theta)\pi(\theta)\,\mathrm{d}\theta$$

depends on the data D only through the statistic s; *i.e.*, for every prior π, the posterior can be written as $\pi(\theta|s)$.

Intuitively, knowing s we are as informed about the parameter θ as when we know all the data collected.

There may be several sufficient statistics available for estimating a parameter. Clearly, from among these we would prefer to make use of one of lowest dimensionality. For a given parameter, there actually may be more than one sufficient statistic of lowest dimensionality and it may be vector valued.

An easy way to find a sufficient statistic is to examine the likelihood for the kind of factorization displayed in

1'5. *Lemma.* – Let the likelihood $L(\theta|D)$ factor so that

$$L(\theta|D) = g(s|\theta)\,h(D)\,,$$

where h does not depend on θ. Then s is a *sufficient statistic* for θ.

Proof. For an arbitrary prior density π, the corresponding posterior density is given by

$$\pi(\theta|D) = L(\theta|D)\pi(\theta)\Big/\int_{\Theta} L(\theta|D)\pi(\theta)\,\mathrm{d}\theta =$$

$$= g(s|\theta)\,h(D)\pi(\theta)\Big/\int_{\Theta} g(s|\theta)\,h(D)\pi(\theta)\,\mathrm{d}\theta = g(s|\theta)\,\pi(\theta)\Big/\int_{\Theta} g(s|\theta)\pi(\theta)\,\mathrm{d}\theta\,.$$

The last expression clearly depends on D only through the statistic s. Thus, by definition 1.4, s is sufficient for θ. ||

For example, if $x_1, \ldots, x_n$ are independent lifelengths in the exponential model, given λ, then n and $s \overset{\text{def}}{=} \sum_1^n x_i$ is a sufficient statistic for the failure rate λ since

$$L(\lambda|x_1, \ldots, x_n) = \lambda^n \exp\left[-\lambda \sum_1^n x_i\right] = \lambda^n \exp[-\lambda s]\,.$$

The sample space. The *sample space* is the space or set of possible sample outcomes. If we observe the lifetimes of n units, the sample space is

$$S = \{(x_1, x_2, \ldots, x_n)|x_i \geqslant 0, 1 \leqslant i \leqslant n\}\,. \tag{1.7}$$

For $n = 2$, S is the positive quadrant (see fig. 1.2).

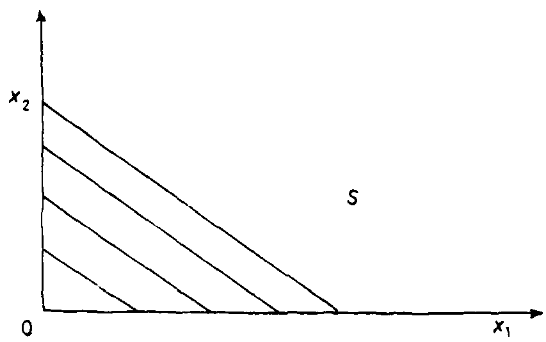


Fig. 1.2.

However, we may just as well consider another sample space. Suppose we are told only the ordered lifetimes

$$x_{(1)} \leqslant x_{(2)} \leqslant \ldots \leqslant x_{(n)}\,,$$

i.e. we no longer know which unit fails at time $x_{(i)}$, $1 \leqslant i \leqslant n$. The sample space corresponding to the ordered lifetimes is now

$$S_0 = \{(x_{(1)}, x_{(2)}, \ldots, x_{(n)})|0 \leqslant x_{(1)} \leqslant \ldots \leqslant x_{(n)}\}\,. \tag{1.8}$$

For $n = 2$, we now have fig. 1.3.

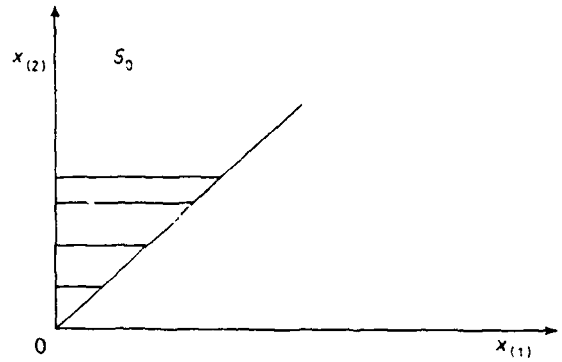


Fig. 1.3.

For sample space (1.7) and the exponential model we have the joint-probability density

$$p(x_1, x_2, \ldots, x_n|\lambda) = \begin{cases} \lambda^n \exp\left[-\lambda \sum_1^n x_i\right], & x_i \geqslant 0,\ 1 \leqslant i \leqslant n, \\ 0, & \text{otherwise}. \end{cases} \tag{1.9}$$

For this case $L(\lambda|x_1, x_2, \ldots, x_n) = p(x_1, x_2, \ldots, x_n|\lambda)$.

On the other hand, for sample space (1.8) and the exponential model we have

$$p_0(x_{(1)}, \ldots, x_{(n)}|\lambda) = \begin{cases} n!\,\lambda^n \exp\left[-\lambda \sum_1^n x_{(i)}\right], & 0 \leqslant x_{(1)} \leqslant \ldots \leqslant x_{(n)}, \\ 0, & \text{otherwise}. \end{cases} \tag{1.10}$$

The factor $n!$ in (1.10) follows from the fact that the ordered observations can result from any one of $n!$ permutations of the observations $x_1, x_2, \ldots, x_n$. For this case

$$L_0(\lambda|x_{(1)}, \ldots, x_{(n)}) = p_0(x_{(1)}, \ldots, x_{(n)}|\lambda),$$

where L_0 is the likelihood.

From (1.9) and (1.10) we see that

$$L(\lambda|x_1, x_2, \ldots, x_n) \propto L_0(\lambda|x_{(1)}, x_{(2)}, \ldots, x_{(n)}). \tag{1.11}$$

It follows that, for any prior for $\tilde{\lambda}$, the posterior density for $\tilde{\lambda}$ will be the same no matter which of the above sample spaces we choose. From (1.11) and lemma 1.5 we see that the order statistics $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ are sufficient for λ.

2. – Selected life test sampling plans.

In this section we illustrate the application of the concepts and methods of sect. 1 when estimating under each of several commonly used life test sampling plans.

Sampling plan (a). Complete observation until a specified number of failures have occurred. A popular plan consists of putting n items on test and observing the failure times of the first k failures, where k is specified in advance. The motivation for following this plan is to save time in determining an estimate of the exponential failure rate.

Let

$$x_{(1)} \leqslant x_{(2)} \leqslant \ldots \leqslant x_{(k)}, \qquad 1 \leqslant k \leqslant n,$$

denote the successive times of failure of the earliest k failures; $x_{(1)}, \ldots, x_{(k)}$ are called the first k *order statistics* in a sample of size n. The likelihood of this observed outcome under the exponential model is given by

$$L(\lambda|D) = \frac{n!}{1!\ldots1!(n-k)!}\left[\prod_{i=1}^{k} \lambda \exp[-\lambda x_{(i)}]\right] \exp[-(n-k)\lambda x_{(k)}].$$

(We follow the usual convention that $0! \equiv 1$.)

Simplifying, we have

$$L(\lambda|D) = \frac{n!}{(n-k)!}\lambda^k \exp\left[-\lambda\left[\sum_{i=1}^{k} x_{(i)} + (n-k)x_{(k)}\right]\right]. \tag{2.1}$$

To verify the expression preceding (2.1), note that the combinatorial coefficient $n!/1!\ldots1!(n-k)!$ represents the number of ways of choosing one observation to correspond to each of $x_{(1)}, x_{(2)}, \ldots, x_{(k)}$ and $n-k$ observations for the $n-k$ unobserved failure times, from among the n failure times (of which only the first k are actually observed). The product factor represents the joint density of the k actually observed failure times, given λ. Finally, the last factor represents the probability that $n-k$ lifelengths each exceed $x_{(k)}$, given λ.

The expression in the exponent of (2.1)

$$\sum_{1}^{k} x_{(i)} + (n-k)x_{(k)} = nx_{(1)} + (n-1)(x_{(2)} - x_{(1)}) + \ldots \\ \ldots + (n-k+1)(x_{(k)} - x_{(k-1)}) \stackrel{\text{def}}{=} T \tag{2.2}$$

represents the *total time on test until the k-th failure.* Note that it is comprised of $nx_{(1)}$, the total time on test observed until the first failure, of $(n-1)(x_{(2)} - x_{(1)})$, the total time on test observed between the first and second failures, ..., and of $(n-k+1)(x_{(k)} - x_{(k-1)})$, the total time on test observed between the penultimate and the last observed failures. (Of course, it is understood that, after an item fails, it is no longer under observation.)

The total time on test statistic turns out to be a very important and useful statistic not only in the exponential model, but also, after appropriate generalization, in a large number of other models involving incomplete data. In the exponential model, the total time on test and the number of observed failures constitute a sufficient statistic for λ. In the presently considered sampling plan under which k, the number of observed failures, is specified in advance, we see from (2.1) that (k, T) *is sufficient for* λ.

Suppose we assume a gamma prior on λ

$$\pi(\lambda) = b^a \lambda^{a-1} \exp[-b\lambda]/\Gamma(a) .$$

Using (2.1), we obtain for the posterior

$$\pi(\lambda|D) = [b + T(x_{(k)})]^{k+a} \lambda^{k+a-1} \exp[-\lambda[b+T]]/\Gamma(k+a) , \tag{2.3}$$

also a gamma, but with the shape parameter a of the prior density replaced by $a+k$ and the scale parameter b of the prior density replaced by $b + T(x_{(k)})$. Note that the increment in the scale parameter is $T(x_{(k)})$, the observed total time on test. The mode of the posterior density is

$$\lambda^0 = (k + a - 1)/[b + T] .$$

It is interesting to note that, for $\pi(\lambda) \equiv c$, the mode of the posterior is exactly the well-known MLE:

$$\hat{\lambda} = k/T .$$

However, it should be emphasized that this prior is *improper* in the sense that $\int_0^\infty \pi(\lambda)\, d\lambda = \infty$.

We would expect that, after collecting a set of data from the exponential distribution, we would have more information concerning the unknown parameter λ than before; more precisely, the peakedness of the density of $\tilde{\lambda}$ might increase or the coefficient of variation might decrease. The *coefficient of variation* of a distribution is the ratio of the standard deviation (assumed finite) to the absolute value of the mean (assumed nonzero). In the case of the gamma prior density, the mean is a/b, the variance is a/b^2, and so the corresponding coefficient of variation is $a^{-\frac{1}{2}}$.

Under the present sampling plan, the posterior distribution, given in (2.3), is also a gamma distribution, but with the shape parameter (the updated « a » value) now $a + k$. It follows that the coefficient of variation is now reduced to $(a+k)^{-\frac{1}{2}}$. Thus, for fixed a, the coefficient of variation decreases roughly as $k^{-\frac{1}{2}}$, where k denotes the observed number of failures. This simple calculation gives us a quantitative notion as to our relative uncertainty concerning λ both before and after observation and, therefore, concerning the decrease in our uncertainty.

Sampling plan (*b*). *Observation terminated at fixed time* (*truncated sampling*). Suppose n units are put on life test at time $t = 0$ and each is observed until failure or fixed time t_0, whichever occurs first. Given the random outcome $K = k$ $(0 \leqslant k \leqslant n)$ observed failures in $[0, t_0]$ and the times of failures, the corresponding likelihood is given by

$$(2.4) \qquad L(\lambda|D) = \frac{n!}{1!\dots 1!(n-k)!}\left[\prod_{i=1}^{k} \lambda \exp\left[-\lambda x_{(i)}\right]\right] \exp\left[-\lambda(n-k)t_0\right],$$

where the product in square brackets is defined as 1 for $k = 0$. The verification of the likelihood expression (2.4) is similar to that obtained under the previous sampling plan, leading to the expression preceding (2.1). One key difference is that the number $n - k$ surviving fixed time t_0 under the present plan is random, while the number $n - k$ surviving past the k-th failure under the earlier plan is specified in advance.

We may rewrite the likelihood $L(\lambda|D)$ as

$$(2.5) \qquad L(\lambda|D) = \frac{n!}{(n-k)!}\lambda^k \exp\left[-\lambda\left[\sum_{i=1}^{k} x_{(i)} + (n-k)\,t_0\right]\right].$$

It is clear from (2.5) and lemma 1.5 that *the pair k and the total time on test*

$$T = \sum_{i=1}^{k} x_{(i)} + (n-b)\,t_0$$

are sufficient for λ. Under the present sampling plan, both k and the $T(t_0)$ are observed, and thus together constitute the data D.

From (2.5) the MLE is computed as

$$\hat{\lambda} = k/T(t_0).$$

Under the present sampling plan it is possible to observe $k = 0$ failures, leading to a MLE of 0. Such an estimate is intuitively unsatisfactory; in this situation an analysis based on the posterior distribution of λ is preferable.

The posterior density resulting from a gamma prior (see (1.4)) is, using (2.5),

$$(2.6) \qquad \pi(\lambda|D) = [b+T]^{k+a}\lambda^{k+a-1} \exp\left[-\lambda[b+T]\right]/\Gamma(k+a).$$

Just as in the previous sampling plan, the posterior coefficient of variation is $(a+k)^{-\frac{1}{2}}$, which is smaller than $a^{-\frac{1}{2}}$, the prior coefficient of variation.

Sampling plan (*c*). *Inverse binomial sampling*. A unit is put on test until it fails or reaches a specified age t_0, whichever occurs first. At this time, the unit is replaced by a new unit. This procedure is repeated sequentially until k

(specified in advance) failures are actually observed. The number N of units that have to be tested until k failures are actually observed is, of course, a random variable. This plan may be used if we have only one test chamber and we are able to test at most one unit at a time.

Let $Y = \min(X, t_0)$, where X has exponential density $\lambda \exp[-\lambda x]$, $x \geqslant 0$. Then, conditional on $Y < t_0$ and t_0, the density of Y is given by

$$\lambda \exp[-\lambda y]/(1 - \exp[-\lambda t_0]) \qquad \text{for } 0 \leqslant y \leqslant t_0 .$$

Thus, if the successive failure ages actually observed are denoted by y_1, $y_2, \ldots, y_k$, the corresponding conditional joint density is given by

$$g(y_1, \ldots, y_k | \lambda, y_i < t_0, 1 \leqslant i \leqslant k) = \prod_{i=1}^{k} \frac{\lambda \exp[-\lambda y_i]}{1 - \exp[-\lambda t_0]} = \frac{\lambda^k \exp\left[-\lambda \sum_1^k y_i\right]}{(1 - \exp[-\lambda t_0])^k} .$$

The probability that $N = n$ units have to be tested in order to observe k actual failures is given by

$$P[N = n | \lambda] = \binom{n-1}{k-1} (1 - \exp[-\lambda t_0])^k \exp[-\lambda t_0 (n-k)] \quad \text{for } n \geqslant k .$$

It follows that, given $N = n$ and observed failure ages $y_1, y_2, \ldots, y_k$, the corresponding likelihood is

$$L(\lambda | D) = \binom{n-1}{k-1} \lambda^k \exp\left[-\lambda \sum_1^k y_i\right] \exp[-\lambda t_0 (n-k)] ,$$
$$0 < y_i \leqslant t_0 , \quad i = 1, \ldots, k .$$

Combining exponentials we obtain

$$(2.7) \qquad L(\lambda | D) = \binom{n-1}{k-1} \lambda^k \exp\left[-\lambda \left[\sum_1^k y_i + (n-k) t_0\right]\right] ,$$
$$0 < y_i \leqslant t_0 , \quad i = 1, \ldots, k .$$

From lemma 1.5, we conclude that (k, T), where

$$T = \sum_{i=1}^{k} y_i + (n-k) t_0 ,$$

is sufficient for λ, since k is fixed in advance.

From (2.7) we also obtain the posterior density for λ based on a gamma prior density as

$$(2.8) \qquad \pi(\lambda | D) = (b + T)^{k+a} \lambda^{k+a-1} \exp[-\lambda(b+T)]/\Gamma(k+a) .$$

Note that the posterior densities are identical under the three sampling plans so far considered! (Compare (2.8), (2.6) and (2.3).)

Remarks. In the three sampling plans considered so far, the number k of observed failures and the total time on test T are all that we need from the observations in order to complete our data analysis; the *sufficiency* of k and T makes this true. Note that this fact holds for any choice of a prior density.

The « stopping rule » in each of the three test plans considered gives no information about the parameter λ, *i.e.* is noninformative about λ. (As the name implies, the *stopping rule* is simply the rule for determining when testing is to stop. The stopping rule is not necessarily the same as the *stopping time.* For example, in the sampling plan (*a*) we test until k failures are observed and then stop further observation.) If the stopping rule were to give information about the parameter λ, then the total time on test T and the observed number of failures would *not* be sufficient for λ.

Finally, note that in the test plans considered thus far, the MLE is the ratio of the number of observed failures to the total time on test. This simple formula for the MLE holds in most of the testing procedures followed under the exponential model.

3. – Inference based on mean life.

Thus far we have discussed inference for the exponential distribution based on the failure rate λ. For many analysts, the mean life θ of the exponential distribution may seem to be the more appropriate parameter to estimate. Note that either parameter determines the other in the exponential model since

$$\theta = \int_0^\infty x\lambda \exp[-\lambda x]\,\mathrm{d}x = \lambda^{-1}.$$

Suppose we consider the simplest type of testing plan: n units are put on test and observed until each fails. The corresponding mutually independent lifelengths given λ are $x_1, x_2, \ldots, x_n$, constituting a complete sample from the exponential density

$$f(x|\theta) = \theta^{-1} \exp[-x/\theta].$$

The *likelihood* of the outcome is given by

$$L(\theta|x_1, \ldots, x_n) = \theta^{-n} \exp\left[-\theta^{-1}\sum_1^n x_i\right]. \tag{3.1}$$

Clearly, the MLE of θ is given by

$$\hat{\theta} = n^{-1} \sum_{1}^{n} x_i .$$

(Note that $\hat{\theta} = (\hat{\lambda})^{-1}$ (see (1.2)).)

Suppose now we have very little prior information on the parameter θ. We, therefore, assume a *rectangular prior density* on $\tilde{\theta}$:

$$\pi(\theta) = M^{-1} \qquad \text{for } 0 \leqslant \theta \leqslant M,$$

where M is large (*). The corresponding *posterior density* for $\tilde{\theta}$ may be computed approximately, by Bayes' theorem (theorem 1.2), as

$$\pi(\theta | x_1, \ldots, x_n) = b^a \theta^{-(a+1)} \exp[-b/\theta]/\Gamma(a) \tag{3.2}$$

for θ, a, $b > 0$, where now $a = n - 1$ and $b = \sum_{1}^{n} x_i$.

The density of (3.2), denoted by $\pi_{a,b}(\theta)$, is called the *inverted gamma density*, since, if $\tilde{\theta}$ is a random variable with density $\pi_{a,b}(\theta)$, then $\tilde{\theta}^{-1} = \tilde{\lambda}$ has gamma density (1.5).

We may verify readily that, if our initial prior is of the form $\pi_{a,b}(\theta)$ of (3.2) and we use any one of the sampling plans, then the corresponding posterior density is also of the form (3.2). However, the parameter a of the prior is replaced by $a + k$ in the posterior, and the parameter b of the prior is replaced by $b + T$ in the posterior. As before, k denotes the number of observed failures and T is the total time on test. This follows readily from the likelihood expression (3.1) and the fact that in this case $T = \sum_{1}^{n} x_i$ since all lifelengths are observed.

The mean of the posterior density. The mean of the inverted gamma density given in (3.2) is readily calculated:

$$\int_0^\infty \theta \{ b^a \theta^{-(a+1)} \exp[-b/\theta]/\Gamma(a) \} \, d\theta = b/(a-1) .$$

For the inverted gamma posterior density in which the parameters are $a + k$

(*) It would be inconsistent mathematically to assume a rectangular prior on *both* $\tilde{\lambda}$ and $\tilde{\theta} = \tilde{\lambda}^{-1}$, even though we have very little prior information on both parameters. We assume a rectangular prior for $\tilde{\theta}$ here to motivate the use of a natural conjugate prior.

(in place of a) and $b + T$ (in place of b), the corresponding mean takes the form

$$\frac{b+T}{a+k-1} = (1-w)\frac{b}{a-1} + w\frac{T}{k}, \tag{3.3}$$

where $w = k/(k + a - 1)$. Thus *the mean of the posterior density may be written as a convex combination of the prior mean $b/(a-1)$ and the maximum-likelihood estimate T/k of the exponential life distribution mean.* Note that, as k, the number of observed failures, increases, the posterior mean attaches more weight to the MLE of the true mean and less weight to the prior mean.

Table 3.I summarizes the properties of the natural conjugate prior density and of the corresponding posterior density for the two different parametrizations of the exponential model.

TABLE 3.I. – *Comparison of alternative parametrizations of the exponential model.*

Parameter	Failure rate, λ	Mean life, θ
likelihood	$\lambda^k \exp[-\lambda T]$	$\theta^{-k} \exp[-T/\theta]$
natural conjugate prior	$b^a \lambda^{a-1} \exp[-b\lambda]/\Gamma(a)$, $a, b > 0$ (gamma)	$b^a \theta^{-(a+1)} \exp[-b/\theta]/\Gamma(a)$, $a, b > 0$ (inverted gamma)
prior mode	$\frac{a-1}{b}$	$\frac{b}{a+1}$
prior mean	a/b	$b/(a-1)$, $a > 1$
prior variance	a/b^2	$b^2/(a-1)^2(a-2)$, $a > 2$
prior coefficient of variation	$a^{-\frac{1}{2}}$	$(a-2)^{-\frac{1}{2}}$, $a > 2$
posterior mean	$(a+k)/(b+T)$	$(b+T)/(a+k-1)$, $a > 1$
posterior variance	$\frac{a+k}{(b+T)^2}$	$\frac{(b+T)^2}{(a+k-1)^2(a+k-2)}$, $a > 2$
posterior coefficient of variation	$(a+k)^{-\frac{1}{2}}$	$(a+k-2)^{-\frac{1}{2}}$, $a > 2$

Note: k = number of observed failures, T = total time on test.

Arbitrary prior density. In our analysis up to now, we have focused mostly on the case in which the prior density is the *natural conjugate prior*. In this subsection, we expand our consideration to cases in which the prior is not necessarily the natural conjugate prior. We obtain results similar to those holding in the natural-conjugate-prior case.

Let $\pi(\theta)$ be a prior density on Θ such that $\{\theta|\pi(\theta) > 0\}$ is an interval on $[0, \infty)$. We show in the appendix that

$$P[\tilde{\theta} > \theta_0|k, T] = \int_{\theta_0}^{\infty} \pi(\theta|k, T)\,d\theta \tag{3.4}$$

is ***decreasing*** (*) ***in*** $k > 0$ for fixed T and ***increasing*** (*) ***in*** T for fixed k, *i.e.* the posterior random variable $\tilde{\theta}$ is ***stochastically decreasing in*** k and ***stochastically increasing in*** T. In particular,

$$E[\tilde{\theta}|k = 0, T] \geqslant E[\tilde{\theta}|k = 0, T = 0];$$

the lower bound is, of course, the mean of the prior density. Thus observing total time on test *without* observing failures tends to change our belief about θ as compared with our prior belief; we tend to believe in a larger θ. However, for the ***natural conjugate*** prior, the variance of $\tilde{\theta}$ given k and T decreases with k but ***increases*** with T. Also the coefficient of variation is ***constant in*** T but ***decreases with*** k (see table 3.I). Hence, if $k = 0$, large values of T tend to make us optimistic regarding θ. However, failures are needed to sharpen the posterior

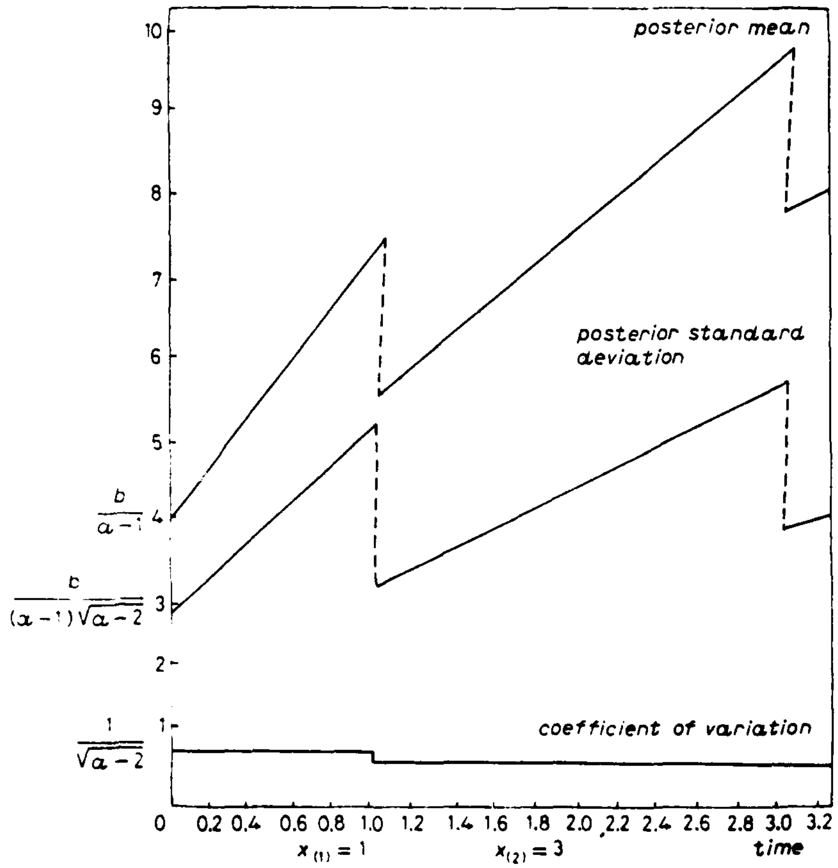


Fig. 3.1. – Posterior mean, posterior standard deviation and posterior coefficient of variation as a function of elapsed test time.

(*) ***Terminology***: Throughout we use the term « increasing » in place of « nondecreasing » and « decreasing » in place of « nonincreasing ».

density. Similarly, for $\tilde{\lambda} = \tilde{\theta}^{-1}$, we can show that for a general prior density on $\tilde{\lambda}$ under mild regularity conditions

$$P[\tilde{\lambda} > \lambda_0 | k, T]$$

increases in k for fixed T and *decreases in T* for fixed k, just as we would expect.

3·1. *Example.* – Suppose our prior density on $\tilde{\theta}$ is the natural conjugate prior with $a = 4$ and $b = 12$. We put 10 units on test. The earliest failure occurs at $X_{(1)} = 1$, followed by a second failure at $X_{(2)} = 3$. In fig. 3.1, we plot the posterior mean, posterior standard deviation and posterior coefficient of variation as a function of t, the test time elapsed. Table 3.I may be used to generate the plots. Note that failures cause vertical drops in the graphs.

In fig. 3.2 we have plotted the posterior density for $\tilde{\theta}$ at selected times during the life test. The posterior density for $t = 0$ is, of course, the prior density. Notice the shape of the posterior density at $t = 1^-$ (*i.e.* just before the first observed failure) and at $t = 1$ (*i.e.* just after the first failure).

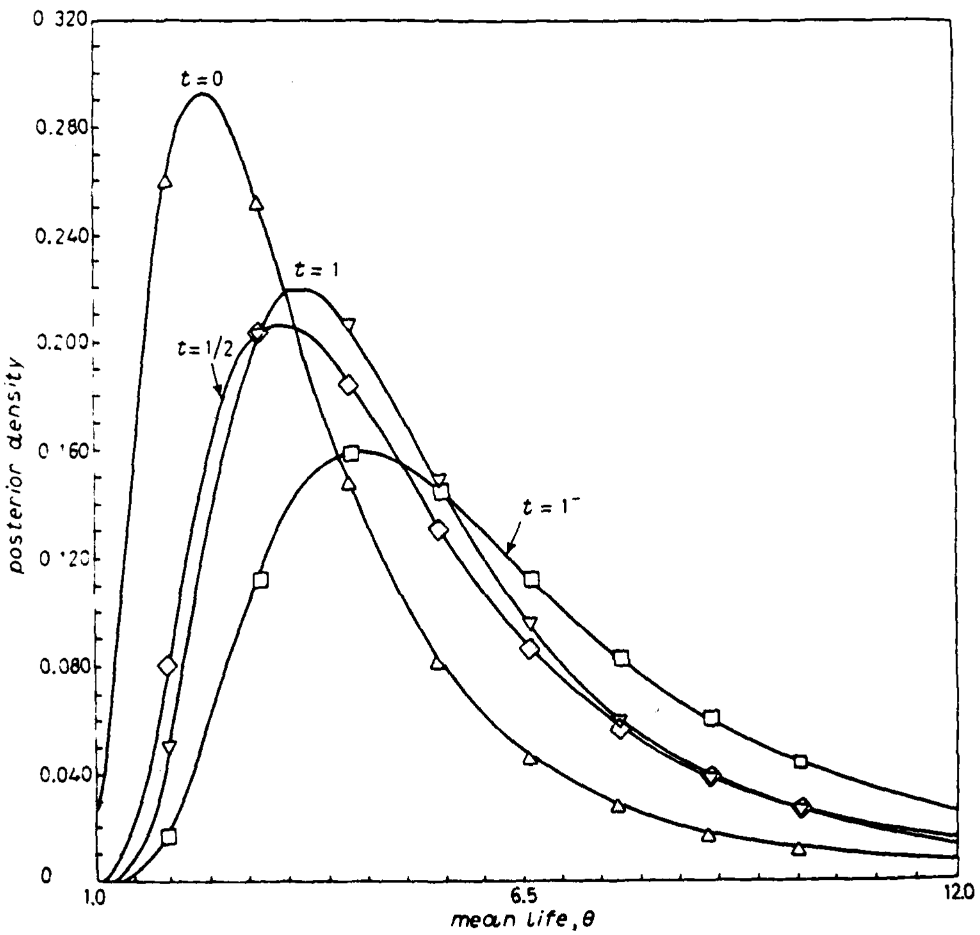


Fig. 3.2. – Posterior densities at selected test times (t).

4. – Notes and references.

Section 1. Books which emphasize Bayesian concepts as well as their applications are [5, 6]. In a series of papers, BASU [4, 7, 8] points out the inadequacy of other approaches to statistical inference. Definition 1.4 of sufficiency is attributed by BASU to KOLMOGOROV [9]. The connection between sample theory (Fisher) sufficiency and Kolmogorov sufficiency is discussed in [10]. Fisher sufficiency implies Kolmogorov sufficiency. The converse is false in general. Natural conjugate priors were introduced by RAIFFA and SCHLAIFER [11]. For a rigorous characterization of natural conjugate priors, see [12]. The reviews of Bayesian statistics by LINDLEY [3, 13] present excellent summaries of recent advances in the subject.

Section 2. EPSTEIN and SOBEL [14, 15] were the first to investigate the properties of the exponential model applied to life test plans. In a series of papers they made an intensive and extensive study of the statistical features of a variety of exponential procedures for life testing. Their work greatly influenced reliability theory and practice at the time and still strongly influences current statistical practice in reliability. Government Handbook H-108 and its subsequent modifications represent the Government's «seal of approval» and its effort to implement the theory by making readily available tables and graphs for easy use of the exponential model [16].

Section 3. Stochastic monotonicity properties of the posterior mean are derived using the concepts and methods of total positivity. (A comprehensive and authoritative treatment of total positivity may be found in [17].) Theorem A.1 in the appendix is similar to lemma 1, p. 276, of Karlin and Rubin [18].

In the mathematical insurance literature formula (3.3) is called the credibility formula. JEWELL [19, 20] has discussed Bayesian life testing.

* * *

This research was supported by the Air Force Office of Scientific Research (AFSC), USAF, under Grant AFOSR-77-3179 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

APPENDIX

Posterior distributions corresponding to arbitrary priors and likelihoods with the monotone ratio property.

We will prove the results mentioned in sect. 3.

A.1 *Theorem.* Let $\pi_0(\theta)$ be a prior density on Θ. Let data $D = (k, T)$ and likelihood $L(\theta|k, T)$ be given such that for $\theta_1 < \theta_2$

$$L(\theta_1|k, T)/L(\theta_2|k, T)$$

is increasing in k and decreasing in T. Let $g(\theta)$ be increasing in θ. Then

$$\int_{\Theta} g(\theta)\pi(\theta|k, T)\,\mathrm{d}\theta \tag{A.1}$$

is decreasing in k (T fixed) and increasing in T (k fixed). The proof will be given shortly.

A.2 *Corollary.* If $L(\theta|k, T) \propto \theta^{-k}\exp[-T/\theta]$, then

$$P[\tilde{\theta} > \theta_0|k, T] = \int_{\theta_0}^{\infty} \pi(\theta|k, T)\,\mathrm{d}\theta \tag{A.2}$$

is decreasing in k (T fixed) and increasing in T (k fixed). (See sect. **3** and (3.4).)

Proof. In theorem A.1, let $g(\theta) \equiv 1$ for $\theta \geqslant \theta_0$ and 0 otherwise. ‖

A.3 *Corollary.* Let $L(\lambda|k, t) \propto \lambda^k \exp[-\lambda T]$ and $g(\lambda)$ be increasing in λ. Then

$$\int_0^{\infty} g(\lambda)\pi(\lambda|k, T)\,\mathrm{d}\lambda \tag{A.3}$$

is increasing in k (T fixed) and decreasing in T (k fixed).

Clearly (A.1) ((A.3)) implies that all moments of the posterior distribution are decreasing (increasing) in k and increasing (decreasing) in T.

Proof of Theorem A.1. For $\theta_1 < \theta_2$,

$$L(\theta_1|k, T)/L(\theta_2|k, T)$$

increasing in k implies that for $k_1 < k_2$

$$L(\theta|k_1, T)/L(\theta|k_2, T)$$

is increasing in θ. It is easy to see that this in turn implies that

$$\pi(\theta|k_1, T)/\pi(\theta|k_2, T) \tag{A.4}$$

is also increasing in θ. Let

$$A = \{\theta|\pi(\theta|k_1, T) > \pi(\theta|k_2, T)\}$$

and

$$B = \{\theta|\pi(\theta|k_1, T) < \pi(\theta|k_2, T)\}\,.$$

Let $a = \inf_{\theta \in A} g(\theta)$ and $b = \sup_{\theta \in B} g(\theta)$. Then the ratio in (A.4) is increasing in θ implies that $a \geqslant b$.

Now

$$\int_{\Theta} g(\theta)[\pi(\theta|k_1, T) - \pi(\theta|k_2, T)]\,\mathrm{d}\theta \geqslant a\int_{A}[\pi(\theta|k_1, T) - \pi(\theta|k_2, T)]\,\mathrm{d}\theta +$$

$$+\, b\int_{B}[\pi(\theta|k_1, T) - \pi(\theta|k_2, T)]\,\mathrm{d}\theta = (a-b)\int_{A}[\pi(\theta|k_1, T) - \pi(\theta|k_2, T)]\,\mathrm{d}\theta \geqslant 0\,.$$

Hence, the integral in (A.1) is decreasing in k.

A similar argument proves that the integral in (A.1) is increasing in T. ‖

REFERENCES

[1] R. E. BARLOW, R. H. TOLAND and T. FREEMAN: *Stress-Rupture Life of Kevlar/Epoxy Spherical Pressure Vessels*, UCID-17755, Part 3, Lawrence Livermore Laboratories, Livermore, Cal. (1979).

[2] A. W. MARSHALL and F. PROSCHAN: *Inconsistency of the Maximum Likelihood Estimator of the Increasing Failure Rate Average Distribution* (in progress).

[3] D. V. LINDLEY: *Bayesian Statistics, A Review* (SIAM, Philadelphia, Penn., 1972).

[4] D. BASU: *Sankhya*, Series A, **37**, Part 1, 1 (1975).

[5] M. H. DEGROOT: *Optimal Statistical Decisions* (McGraw-Hill New, York, N. Y., 1970).

[6] R. L. WINKLER: *An Introduction to Bayesian Inference and Decision* (Holt, Rinehart and Winston, Inc., New York, N. Y., 1972).

[7] D. BASU: *J. Am. Stat. Assoc.*, **72**, 355 (1977).

[8] D. BASU: *J. Stat. Planning Inference*, **2**, 1 (1978).

[9] A. N. KOLMOGOROV: *Izv. Akad. Nauk SSSR, Ser. Mat.*, **6**, 3 (1942), in Russian.

[10] SHIH-CHUAN CHENG: *A mathematical study of sufficiency and adequacy in statistical theory*, Ph.D. Thesis, Florida State University, Tallahassee, Fl. (1978).

[11] H. RAIFFA and R. SCHLAIFER: *Applied Decision Theory*, Graduate School of Business Administration, Harvard University, Boston, Mass. (1961).

[12] R. DIACONIS and D. YLVISAKER: *Ann. Stat.*, **7**, 269 (1979).

[13] D. V. LINDLEY: *Scand. J. Stat.*, **5**, 1 (1978).

[14] B. EPSTEIN and M. SOBEL: *J. Am. Stat. Assoc.*, **48**, 486 (1953).

[15] B. EPSTEIN and M. SOBEL: *Ann. Math. Stat.*, **25**, 373 (1954).

[16] *Handbook* H-108, *Sampling Procedures and Tables for Life and Reliability Testing*, DoD Document.

[17] S. KARLIN: *Total Positivity*, Vol. **1** (Stanford University Press, Stanford, Cal., 1968).

[18] S. KARLIN and H. RUBIN: *Ann. Math. Stat.*, **27**, 272 (1956).

[19] W. S. JEWELL: *Bayesian life testing using the total Q on test*, in *The Theory and Applications of Reliability*, Vol. **1**, edited by C. P. TSOKOS and I. N. SHIMI (Academic Press, New York, N. Y., 1977), p. 49.

[20] W. S. JEWELL: *Bayesian Learn While Waiting, The ASTIN Bulletin*, Vol. **10**, Part 2 (1978), p. 163.

Reprinted From
Theory of Reliability
© 1985, XCIV Corso
Soc. Italiana di Fisica - Bologna - Italy

ATE
LMED
8